Stream Assessments in the Georgia Basin Using the Reference Condition Approach for Benthic Invertebrate Monitoring

Stephanie Sylvestre *and* ***Taina Tuominen***
Environment Canada, Pacific and Yukon Region

Trefor Reynoldson
National Water Research Institute

Abstract
In 1998-2002, more than 50 reference sites from the lower Fraser Valley and eastern Vancouver Island were sampled and added to the existing Fraser River Basin benthic invertebrate database for reference condition. The Fraser River Basin predictive model was re-developed specifically to include these and other sites in the Georgia Basin. More than 40 sites were sampled from streams exposed to agricultural and urban activities in the Georgia Basin and assessed using the reference condition approach based on abundance predictions as well as a RIVPACS-type approach based on taxonomic predictions. Eight sites were also re-sampled in subsequent years to assess repeatability of the assessment. Most assessments were similar in multiple years. The streams indicating the most stress on the invertebrate community were those exposed to urban activities; this stress may be a consequence of dramatic hydrological changes during rain events.

Introduction
Benthic macroinvertebrates are the most commonly used aquatic organisms for bioassessment (Rosenberg and Resh 1993). They are ubiquitous, relatively sedentary and they reflect site-specific conditions. They occur in high diversity so they respond to a variety of environmental stresses. In addition, they are easy to collect and inexpensive to analyses compared with water quality samples. Because of these attributes, these organisms have been used in biomonitoring programs employing both multi-metric and multivariate approaches.

A biomonitoring program based on the reference condition approach was developed in 1994-1998 for the Fraser River Basin in British Columbia from the upper reaches of the Fraser River to the town of Chilliwack, approximately 125 km east of Vancouver (Rosenberg et al. 1999; Reynoldson *et al*. 1997). The program was based on the development of a large database of benthic macroinvertebrate assemblages from minimally disturbed sites throughout the basin encompassing a wide range of habitats. Using multivariate techniques, bioassessment models were developed from this database that predicted the expected invertebrate assemblage at a particular site with specific habitat characteristics. Such models were used to assess sites within the Fraser River Basin suspected to be impacted by various human activities (Reynoldson *et al*. 2001).

In British Columbia, most of the human activities affecting streams occur in the higher population centres in the Georgia Basin. This includes the lower Fraser Valley downstream of Hope and Eastern Vancouver Island. Many of the streams in the Georgia Basin are outside the landscape or geographic range (i.e. ecoregion, latitude, longitude) observed in the Fraser Basin reference database and they have different habitat characteristics from those sampled in the Fraser Basin. These Georgia Basin streams tend to be slower streams with soft bottoms and relatively deep channels compared with the cobble-riffles sampled in the Fraser Basin. The geographical barrier of the Strait of Georgia was also a consideration for the application of the Fraser Basin model. The invertebrate communities on Vancouver Island had not been compared previously to the mainland communities. The objective of this study was to expand the Fraser Basin reference database to build a model that can be used to assess sites in the Georgia Basin exposed to urban and agricultural activities.

Methods

Sampling design
Potential reference areas in the lower Fraser Valley were identified on topographic maps and confirmed by aerial and land reconnaissance. Potential reference areas on Eastern Vancouver Island were identified by local experts and confirmed by land reconnaissance. Sites were selected on a variety of stream orders within the eastern Vancouver Island and the Lower Mainland ecoregions on topographic maps and exact locations were determined by access. Thirty-four reference sites were sampled in the lower Fraser Valley between 1998 and 2000 and 21 reference sites were sampled on Vancouver Island in 2001. Thirty-five sites suspected to be affected by agricultural and urban activities were sampled in between 1998 and 2001. More than 40 environmental variables were recorded or measured at each site. The invertebrate

community was sampled using a triangular kick-net for a timed kick of 3 minutes. The invertebrate samples were identified to the lowest possible taxonomic level. The biological and habitat sampling procedures followed those defined in the biomonitoring program for the Fraser River Basin (Rosenberg *et al.* 1999).

Model Development

The reference condition approach was used to assess streams in the Georgia Basin as it was used in the Fraser River Basin. The rationale and detailed analytical methods are described in Rosenberg *et al.* (1999) and Reynoldson *et al.* (2001). This approach required that a wide range of reference sites be sampled to capture as much variability as possible from which models were developed that link habitat variables to the biological community. Fifty-five reference sites were sampled in the Georgia Basin and were added to 199-site reference database for the Fraser Basin. The Bray-Curtis association matrix was used to describe the communities and cluster analysis using agglomerative hierarchical fusion method with unweighted pair group mean averages (UPGMA) was used to classify the reference site communities into sub-groups. Based on results by Reynoldson *et al.* (2001) of a comparison of different taxonomic models, the taxonomic family level was used for the analysis.

Twenty-nine variables were considered for the development of predictive models (Table 1). Stepwise discriminant function analysis (DFA) was performed using SYSTAT 10 (SPSS 2000) to identify habitat variables that best separated sites into the pre-defined reference groups formed by the cluster analysis. This was followed by an iterative process of selecting the optimal combination of predictor variables.

Table 1. Habitat variables measured for the development of a predictive model (modified from Rosenberg *et al.* 1999).

Scale	Variable	units
Landscape	stream order	7 categories (1-7)
	latitude	decimal degrees
	longitude	decimal degrees
	altitude	feet above sea level
	ecoregion	11 categories[1]
Site	flow state	3 categories (riffle, run, pool)
	% macrophyte cover	5 categories (0, 25, 50, 75, 100)
	presence of grasses	2 categories (0, 1)
	presence of shrubs	2 categories (0, 2)
	presence of deciduous tress	2 categories (0, 3)
	presence of coniferous trees	2 categories (0, 4)
	riparian vegetation	10 categories (sum of above 4 variables)
Channel	wetted width	m
	bankfull width	m
	average depth	cm
	maximum depth	cm
	slope	m/m
	average velocity	m/s
	maximum velocity	m/s
	substrate framework	7 categories[2]
	substrate matrix	7 categories[3]
	substrate embeddedness	5 categories[4]
	gravel	% of interstitial material
	sand	% of interstitial material
	silt	% of interstitial material
	clay	% of interstitial material
Water-column	pH	relative units
	alkalinity	mg/L
	conductivity	μS/cm

[1]B.C. Ministry of Environment, Lands and Parks 1991.

[2]Dominant substrate size in sampling area: 1, <0.02cm; 2, 0.2-0.5cm; 3, 0.5-2.5cm; 4, 2.5-5 cm; 5, 5-10 cm; 6, 10-25 cm; 7, >25 cm.

[3]Substrate size surrounding dominant substrate: 1, <0.02cm; 2, 0.2-0.5cm; 3, 0.5-2.5cm; 4, 2.5-5 cm; 5, 5-10 cm; 6, 10-25 cm; 7, >25 cm.

[4]Estimated in sampling area: 1, completely embedded; 2, 75% embedded; 3, 50% embedded; 4, 25% embedded; 5, unembedded.

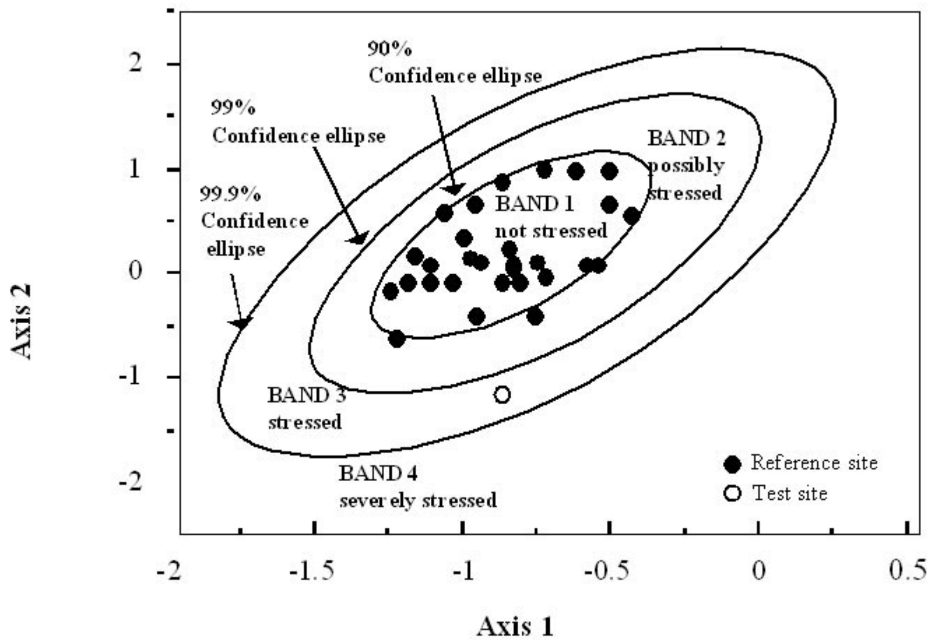

Figure 1. Assessment of a test site in ordination space by hybrid multi-dimensional scaling with the group of reference sites predicted by the DFA model and the bands of biological quality defined by probability ellipses.

The final model is a probabilistic predictive model that assigns probabilities of a test site to belonging to each of the reference groups. The accuracy of the models was evaluated in SYSTAT 10 (SPSS 2000) by re-substitution and cross-validation in the discriminant analysis procedure. Re-substitution uses a reference site to both estimate the model and evaluate the success of the model, whereas cross-validation evaluates the model, one site at a time, with a site that is not used to derive the model. Various models were examined by iteration by adding and replacing variables. The optimal set of variables was defined as that which had the lowest error rate from the cross-validation in DFA.

Assessment of test sites

The reference condition approach assumes that the reference sites represent normal or unimpaired conditions and that the degree of similarity between a test site and the reference sites represents the quality of that site. With the DFA model, a test site is predicted to being similar to each group of reference sites with a specified probability based on the habitat variables at the test site. The reference group to which the test site has the highest probability of belonging represents the community that is expected at the test site. The assumption is that if a test site community is different than what is expected based on the habitat characteristics then there must be some stress exerted on that community. To determine similarity, the test site community is plotted in ordination space with the group of reference site communities to which it is predicted (Rosenberg *et al.* 1999). The variation in the reference sites represents the variation that is expected at the test site. Based on how similar a test site community is to the group of reference sites, categories of biological quality are defined using probability ellipses around the “cloud” of reference sites in ordination space (Figure 1). Three probability ellipses are drawn, 90% probability, 99% probability, and 99.9% probability. The first band includes all sites within the 90% ellipse, which represents an unstressed community or communities that are equivalent to reference. The second band is the area between the 90% and the 99% confidence ellipse. This area is considered potentially stressed or possibly different from reference since some reference sites will also fall into this region. The third band is the area between the 99% and the 99.9% ellipse where a community would be stressed or is different from reference. The fourth band is the area outside of the 99.9% ellipse where a community would be much stressed or very different from reference. The discriminant model was used to assess 35 test sites. The test site was compared with the group of reference sites to which it had the highest probability of belonging based on the habitat features and was assessed using the procedure described above.

Predicting invertebrate assemblages

There is inherent uncertainty in assigning a test site to a reference group. The confidence that the user has in the assessment depends on the level of probability of a test site belonging to the predicted reference group. To address this uncertainty, another way to examine the site is to look at the expected versus observed taxa found at a site by weighting the prediction of the expected taxa at a test site to the probability of the test site belonging to each of the reference groups. We can eliminate the uncertainty associated with predicting a test site to a particular group and combine the results of the model to predict the invertebrate assemblage. By calculating the frequency of each taxon occurring in each reference group and multiplying that by the probability of the test site belonging to each of the groups, we can calculate the probability of each taxon occurring at a test site. This assessment using presence-absence is an approach used in the UK called RIVPACS (River Invertebrate Prediction and Classification System; Wright *et al.* 1984) and a similar approach is used in Australia called AusRivAS (AUStralian RIVer Assessment Scheme, Parson and Norris 1996). Our ordination assessment procedure is heavily weighted on relative abundances, particularly by the dominant taxa, and does not use presence-absence data because we believe that variation in abundance is important in distinguishing impaired sites. Although we believe the assessment should be based on abundance data rather than presence-absence data. Presence-absence data can be useful in understanding the community response and can be used in conjunction with the ordination assessment. Observed taxa to expected taxa ratios were calculated for each test site using taxa that had a probability of occurring of 50% or greater.

Results and Discussion

Model Development

The extended Fraser/Georgia Basins reference database consists of an additional 54 reference sites and 19 taxa. Cluster analysis identified four biological reference groups (Figure 2). A minimum group size of 10 sites was selected and used as the stropping rule for the classification (Rosenberg *et al*. 1999). The classification results are similar to what was observed in the Fraser Basin database (Rosenberg *et al*. 1999). The reference sites selected in the lower Fraser Valley increased the number of sites in the original Fraser River database with deeper channels, soft bottoms and slower flows; however the communities were distributed among the four reference groups (Table 2). The communities did appear to expand the cloud of reference sites in ordination space along axis 1 and axis 2 (Figure 3). Conversely, the Vancouver Island sites were distributed well within the range of communities seen in the Fraser Basin (Figure 3) and were primarily classified to Group 1; however some communities were also classified to Group 3 (Table 2). This implies that at the family level, the communities on the island are not distinctly different from the Fraser Basin communities as a result of the geographic barrier of the Strait of Georgia and we can extend the reference database from the mainland to the island.

Table 2. Distribution of sites from the Fraser River Basin, the Lower Fraser Valley and Vancouver Island among the reference groups defined by the cluster analysis.

	Group 1	Group 2	Group 3	Group 4	**Total Sites**
Fraser Basin	62	11	66	60	**199**
Lower Fraser Valley	7	5	14	8	**34**
Vancouver Island	17	0	4	0	**21**
Total Sites	**86**	**16**	**84**	**68**	**254**

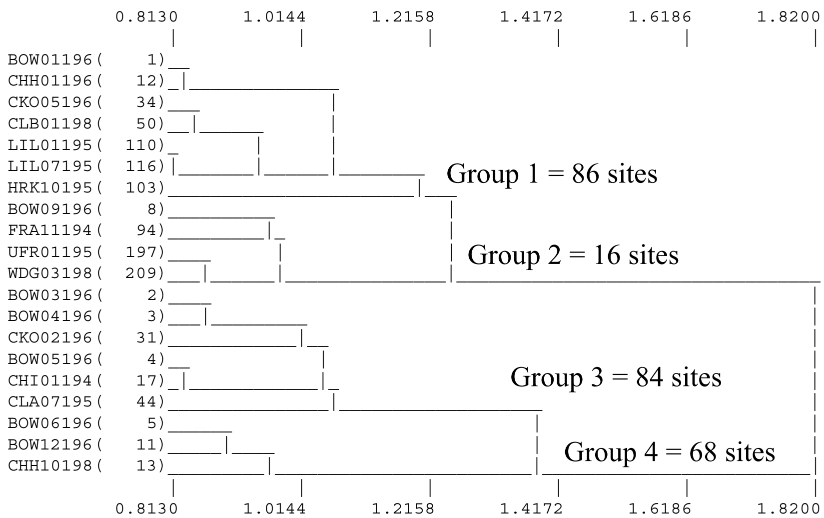

Figure 2. Classification of 254 reference sites based on Bray-Curtis similarity using agglomerative hierarchical fusion, the unweighted pair-group method with arithmetic averages (UPGMA).

The reference sites actually represent a continuum of communities that we have imposed boundaries on through classification into sub-groups (Figure 4). The groups overlap in their biological structure (Table 3) as well as in their habitat features (Table 4). In terms of biological structure, the reference communities differed by both the most common taxa present and the relative abundances of the different taxa. Group 1 and Group 2 sites tended to have overall lower abundances of taxa. Group 2 was the most diverse with common taxa from the Oligochaeta order, which were not commonly found in the other groups. Group 1 was dominated by the Heptageniidae mayfly but also had a wide variety of stoneflies (Plecoptera) and caddisflies (Trichoptera). Trichoptera were not commonly found in Group 2. Groups 3 and 4 were similar in the taxa present and the taxa that were dominant. However, Group 4 had very large abundances of organisms and tended to be heavily dominated by the presence of Chironomidae and Elmidae beetles. Leptophlebiidae mayflies were commonly found in Group 4; these mayflies were not common in the other groups. In terms of habitat features, Group 1 sites tended to be smaller streams with low alkalinity, large substrates and faster velocities. Group 2 sites were generally the Fraser mainstem sites or sites with low gradients, large channels and soft bottoms. Groups 3 and 4 were moderate sized streams that differed primarily in alkalinity, conductivity and substrate size.

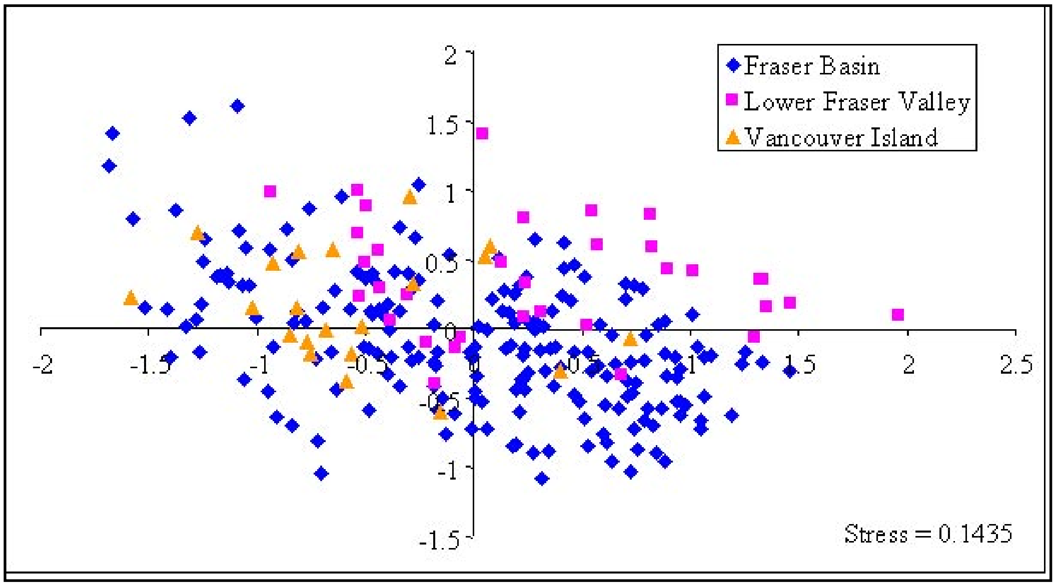

Figure 3. Ordination using hybrid multi-dimensional scaling (HMDS) of 93 families from reference sites in the Fraser River basin, the lower Fraser Valley and Vancouver Island.

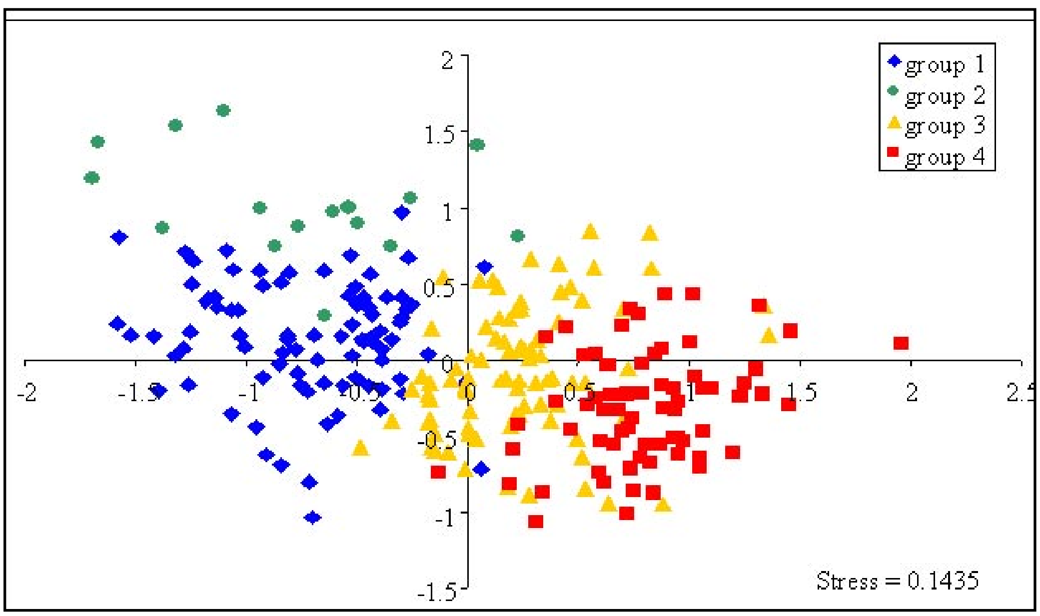

Figure 4. Ordination using hybrid multi-dimensional scaling (HMDS) of 93 families from the four reference groups defined by cluster analysis.

Table 3. Median abundances of the most common taxa in each of the reference groups.

Order	Family	Group 1	Group 2	Group 3	Group 4
	Heptageniidae	117	3	479	1475
	Baetidae	73	3	373	1400
	Ephemerellidae	32	3	229	300
	Leptophlebiidae				150
	Ameletidae		1		
Plecoptera	Nemouridae	13	1	53	300
	Capniidae	4	31	25	150
	Chloroperlidae	25	4	51	100
	Perlodidae	8	1	14	58
	Taeniopterygidae	35			
Trichoptera	Rhyacophilidae	4		4	
	Hydropsychidae	1			
Diptera	Chironomidae	69	46	689	3867
	Empididae	1	2	13	38
	Tipulidae	2	2	16	
Coleoptera	Elmidae				100
Acarina	Limnesiidae		1		
Oligochaeta	Enchytraeidae		2		
	Lumbriculidae		1		
	Naididae		1		
Median Total		**383**	**98**	**1945**	**7938**

Table 4. Selected habitat variables (mean ± SD) for reference groups defined by cluster analysis.

	Group 1 (n=86)	Group 2 (n=16)	Group 3 (n=84)	Group 4 (n=68)
Stream Order	2.6 (1.1)	3.8 (1.7)	2.7 (1.3)	2.8 (1.6)
Alkalinity (mg/L)	35.5 (27.4)	40.8 (29.4)	42.7 (32.0)	58.0 (38.2)
Altitude (fasl)*	2173 (1717)	2059 (1699)	2872 (1753)	3240 (1511)
Bank width (m)	48.7 (54.1)	106.7 (99.2)	31.8 (39.2)	32.4 (60.3)
Average Depth (cm)	28.2 (12.9)	34.5 (22.2)	27.5 (18.5)	21.0 (11.8)
Channel width (m)	18.0 (21.0)	72.1 (69.2)	16.1 (25.1)	15.0 (29.4)
Clay (%)	0.22 (1.5)	0.52 (2.1)	0.11 (1.0)	0.98 (4.4)
Conductivity (uS/cm)	82.5 (69.4)	93.3 (67.4)	82.2 (65.7)	113.9 (79.2)
Embeddedness*	3.5 (1.0)	3.1 (1.2)	3.9 (1.1)	3.8 (1.2)
Framework*	6.7 (1.3)	5.0 (1.9)	5.9 (1.7)	5.0 (1.9)
Gravel (%)	34.9 (32.8)	19.4 (19.2)	32.1 (23.3)	31.2 (26.2)
pH	7.6 (0.6)	7.3 (1.2)	7.5 (0.7)	7.6 (0.7)
Grass in riparian*	0.3 (0.5)	0.3 (0.5)	0.5 (0.5)	0.7 (0.5)
Sand (%)	62.7 (32.3)	76.1 (18.4)	62.6 (26.7)	60.5 (28.1)
Slope	0.024 (0.36)	0.006 (0.013)	0.017 (0.03)	0.009 (0.015)
Average Velocity (m/s)	0.44 (0.20)	0.32 (0.27)	0.37 (0.18)	0.35 (0.19)

*Refer to Table 1 for a description/definition of measurement.

Through discriminant function analysis (DFA) and an iterative process, the optimal predictive model with the lowest error rate consisted of 11 variables. Two combinations of variables resulted in the same error rate, which was similar to the original Fraser Basin model (Table 5). The success of model prediction tends to decline with an increase in the number of sites in the database (Wright 1995) due to the overlap of reference groups and the continuum of biological communities that actually exists. However, an addition of 27% more sites to the Fraser Basin database increased the error rate by only 2% by re-substitution and by 5% by cross-validation for the complete model and by 2% for both re-substitution and cross-validation for the optimal model, suggesting that the predictive performance is relatively robust. Many of the variables were similar in all models suggesting that the predictive variables are also robust. While the scope of this paper examines only the modification of the Fraser Basin model for use in the Georgia Basin, important considerations such as accuracy and precision of the Fraser Basin model have been discussed by Reynoldson *et al*. (1997). In addition Reynoldson *et al*. (2001) also examined sensitivity, robustness, annual variation, seasonal variability, usability, predictive performance and certainty of multivariate and multi-metric models.

Table 5. Predictive performance of Fraser Basin optimal model and redeveloped Fraser/Georgia Basin models using variables selected by discriminant function analysis and iteration.

Model	Variables	Resubstitution % correct	Cross-validation % correct
Fraser Basin			
Complete	all 29 variables	67	53
Optimal (9 variables)	alkalinity, depth(max), channel width, framework, conductivity, ecoregion, grass, latitude, longitude	62	56
Fraser/Georgia Basin			
Complete	all 29 variables	65	48
Optimal #1 (11 variables)	**alkalinity, depth(mean), channel width, framework, conductivity, ecoregion, grass, embeddedness, altitude, bank width, clay**	**60**	**54**
Optimal #2 (11 variables)	alkalinity, depth(mean), channel width, framework, conductivity, ecoregion, grass, embeddedness, latitude, slope, velocity(max)	60	54

Bold indicates the model used for test site assessments.

Assessment of Test Sites

The test sites exposed to a variety of urban and agricultural activities were assessed using the first optimal Fraser/Georgia Basin model indicated in bold in Table 5. None of the sites were predicted to belonging to the Group 2 reference sites. These are the Fraser River mainstem sites with large channels and soft substrates. Despite the soft substrates found in some of the test sites, particularly the agricultural sites, the model predicted with fairly high certainty (Table 6) that they should be similar to reference sites in groups 1, 3, or 4 based on the other habitat features. Only three sites had probabilities of group membership of less than 40% and five sites had probabilities of group membership greater than 80%. The average probability for the 35 test sites was 61%. This indicates that there is some uncertainty of the model prediction for these test sites due to the inherent difficulties in assigning sites to overlapping reference groups. The RIVPACS method of predicting species composition alleviates the problem of misclassification and the analysis of presence-absence data can be used in conjunction with the ordination assessment to address uncertainties.

Table 6. Summary of assessments, the probability of membership to predicted reference group and observed to expected taxa ratios for sites exposed to urban and agricultural activities.

	Agricultural sites		Urban sites		
Assessment	# sites	Range of probabilities of group membership	# sites	Range of probabilities of group membership	O:E ratio range
not stressed	2	38-62%	4	39-67%	0.81-1.27
potentially stressed	7	37-92%	11	43-72%	0.13-1.26
stressed	0		4	62-97%	0.25-0.72
severely stressed	2	51-68%	5	41-71%	0.23-0.54

The RIVPACS method is based on presence-absence data, which uses the weighted probabilities for the four reference groups and the frequency of taxon occurrence in each of four groups. Based on the assumptions of the RIVPACS method, the higher the ratio of observed to expected taxa (O:E ratio) the better the quality of the site. Generally the ratios are in concordance with the ordination assessments, particularly for the very good (not stressed) and the very poor (severely stressed sites). The variance in agreement is greatest for the potentially stressed sites. Where the ordination assessment was poor (stressed) yet the O:E ratio was high, the expected taxa were present but not in the abundances expected based on the comparison to the predicted group of reference sites. This is where we believe that abundance is important in distinguishing an impaired site from a reference site rather than just relying on presence-absence data.

The reverse also occurred at seven sites (four urban sites and three agricultural sites) where the O:E ratio was low but the ordination assessment was only potentially stressed. This is because only a few expected taxa were present, but they were the most dominant taxa in the relative abundances that were expected based on the ordination assessment. This occurred most often for sites predicted to belonging to group 4, where the site is expected to have a large number of organisms dominated by Chironomidae and Beatidae. In most cases, these were the only two of the expected taxa ($p>0.50$) found; but because they were abundant and dominant the ordination method found the site to be of better biological quality than the O:E ratio suggests. For the seven sites that showed this discrepancy, they should be considered stressed communities rather than potentially stressed. It is in this scenario where the absence of the less dominant taxa should be examined using O:E ratios in addition to the ordination assessment for the final interpretation of the biological quality of the site.

Eleven agricultural sites and 24 urban sites were assessed. The urban sites were more often found to have stressed or severely stressed communities based on what is expected from the habitat features than the agricultural sites. These assessments do not address cause and effect; rather they examine cumulative impacts with the predictive capacity to indicate what the community should look like in the absence of an anthropogenic stress.

Future directions

The development of the large reference database and the bio-assessment models for the Fraser/Georgia Basins provides a valuable tool for environmental assessment. These will soon be available online to users interested in bioassessment. This approach is being used elsewhere in Canada—in the Great Lakes and the Atlantic Region—and will soon extend to other parts of the country. It is intended that this approach will be integrated in a national program called Canadian Aquatic Biomonitoring Network. (CABIN), that will be a collaborative effort between Environment Canada and users interested in bio-assessment. CABIN provides, stores and maintains the invertebrate databases where data can be integrated from across the country. CABIN will also provide user-friendly bio-assessment software with built-in multivariate statistics and predictive models so that the user needs only limited knowledge of the detailed analysis. The CABIN website with the standardized sampling protocols and bio-assessment tools is currently under development. For more information regarding CABIN, contact Dr. Trefor Reynoldson (trefor.reynoldson@acadiau.ca) or Stephanie Sylvestre (stephanie.sylvestre@ec.gc.ca).

References

B.C. Ministry of Environment, Lands and Parks, 1991, *Ecoregions of British Columbia.* Habitat Inventory Section, Wildlife Branch, BC. Ministry of Environment, Lands and Parks, Victoria, BC.

Parsons, M. and Norris, R.H. 1996, The effect of habitat-specific sampling on biological assessment of water quality using a predictive model. *Freshwater Biology* **36:**419-434.

Reynoldson, T.B., Norris, R.H., Resh, V.H., Day, K.E., and Rosenberg, D.M. 1997, The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. *Journal of the North American Benthological Society* **16:**833-852.

Reynoldson, T.B., Rosenberg, D.M., and Resh, V.H. 2001, Comparison of models predicting invertebrate assemblages for biomonitoring in the Fraser River catchments, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences* **58**:1395-1410.

Rosenberg, D.M. and Resh, V.H. 1993, *Freshwater biomonitoring and benthic macro-invertebrates*. Chapman and Hall, New York.

Rosenberg, D.M., Reynoldson, T.B., and Resh, V.H. 1999, *Establishing reference condition for benthic invertebrate monitoring in the Fraser River Catchments, British Columbia, Canada*. Aquatic and Atmospheric Sciences Division, Environment Canada, Vancouver, BC. DOE FRAP 1998-32.

SPSS 2000, *SYSTAT 10*. SPSS Inc., Chicago, Illinois.

Wright, J.F. 1995, Development and use of a system for predicting the macro-invertebrate fauna in flowing waters. *Australia Journal of Ecology* **20**:181-197.

Wright, J.F., Moss, D., Armitage, P.D., and Furse, M.T. 1984, A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. *Freshwater Biology* **14**:221-256.